9-2015

Candy Crush Combinatorics

Dana Rowland
Merrimack College, rowlandd@merrimack.edu

Candy Crush Combinatorics

Dana Rowland

Dana Rowland (rowlandd@merrimack.edu) received her B.A. from the University of Notre Dame and an M.S. and Ph.D. from Stanford University. She is co-chair and associate professor of mathematics at Merrimack College in North Andover, Massachusetts. When she is not doing mathematics or playing Candy Crush, Rowland enjoys board games, hiking, soccer, and playing the bassoon.

Walk in to any waiting room and you are likely to find someone passing time playing the game Candy Crush Saga. The game was launched on Facebook in 2012 by the game company King and is played in over 200 countries on various mobile and web platforms [**1**]. Although the game is free, it remains one of the top grossing iPhone and iPad apps. Its addictive nature has captured the interest of students and faculty alike. After seeing a student playing before class, I started thinking about some of the interesting mathematical questions you can ask about Candy Crush. This paper is an attempt to justify the many hours I have spent playing the game.

In Candy Crush Saga, differently colored candies are arranged in a grid. To clear a level, the player must swap adjacent candies in order to match three or more candies of the same color. A valid starting configuration of a game of Candy Crush will not have 3 consecutive candies of the same color in a row or column. Also, it must be possible to swap two adjacent candies to obtain at least 3 consecutive candies of the same color. This leads to two combinatorial questions:

- How many ways can you fill an m by n grid using q colors without three consecutive candies of the same color?
- How many of those grids will contain a move?

We begin by answering these questions when $m = 1$; that is, Candy Crush on a line.

Candy Crush on a line, avoiding "three in a row"

In this section we construct $1 \times n$ grids which avoid having 3 candies in a row of the same color. Note that adjacent pairs of the same color are allowed—a valid configuration might have as many as $\lfloor n/2 \rfloor$ adjacent pairs.

Let a_n be the number of ways to color a line of n candies using q colors without coloring three consecutive candies the same color. It immediately follows that $a_1 = q$ and $a_2 = q^2$. For $n \geq 3$, note that a line of candy either ends with a single candy of a color or pair of candies of the same color. A line of length n ending with a single candy can be obtained in $(q-1)a_{n-1}$ ways and a line of length n ending with a pair

http://dx.doi.org/10.4169/college.math.j.46.4.255
MSC: 05A19

can be obtained in $(n-1)a_{n-2}$ ways. Therefore, a_n, the total number of ways to color a line of n candies without having 3 in a row, is given by the recursive formula

$$a_1 = q,\ a_2 = q^2,\ a_n = (q-1)(a_{n-1} + a_{n-2}). \tag{1}$$

By using the characteristic equation, we obtain the closed form

$$a_n = \frac{q\left(\left(q-1+\sqrt{q^2+2q-3}\right)^{n+1} - \left(q-1-\sqrt{q^2+2q-3}\right)^{n+1}\right)}{2^{n+1}(q-1)\sqrt{q^2+2q-3}}.$$

Alternatively, we can obtain an expression for a_n directly by counting the number of ways to get a line of candies which have at most two of the same color in a row. To determine the number of configurations with exactly k pairs of the same color, for $0 \le k \le n/2$, we split the n positions into $n-k$ slots, where we will fill k of those slots with pairs of candies of the same color and the remaining slots with single candies. See Figure 1.

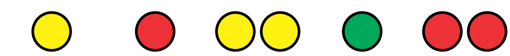

Figure 1. A configuration with $n = 7$ candies, $k = 2$ pairs, $n - k = 5$ slots, and $q = 3$ colors.

There are $\binom{n-k}{k}$ ways to determine which of the slots are filled with pairs. Candies in adjacent slots may not be the same color, so there are q ways we can assign a color to the first slot and $q-1$ ways for each of the remaining slots. Thus, the number of ways to fill in a $1 \times n$ grid avoiding 3 in a row of the same color is given by

$$a_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-k}{k} q(q-1)^{n-k-1}. \tag{2}$$

Candy Crush Saga begins with a 5 by 8 grid in Level 1 with six colors of candy. In that case, (1) becomes $a_n = 5(a_{n-1} + a_{n-2})$, $a_1 = 6$, $a_2 = 36$ and (2) becomes

$$a_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-k}{k} 6 \cdot 5^{n-k-1}.$$

These yield $a_5 = 7{,}200$ and $a_8 = 1{,}444{,}500$, so there are 7,200 different possible starting configurations for a particular column, and similarly there are 1,444,500 different possible starting configurations for a given row in Level 1 of Candy Crush Saga.

A valid line of Candy Crush containing a move

We will say a line of candies is *valid* if it avoids three consecutive candies of the same color and we say it *contains a move* if it is possible to swap adjacent candies in order to get 3 candies in a row. In this section we count how many valid lines of length n contain a move.

Let c, d, and e denote distinct colors. Observe that

- Since a line does not have three consecutive candies of the same color, it must end in one of the patterns ccd, cdc, cdd, or cde. See Figure 2.

Figure 2. Possible ending patterns for a line.

- A line contains a move if and only if at least one of the patterns $ccdc$ or $cdcc$ appears somewhere in the line. See Figure 3.

Figure 3. A line with a move must contain one of these patterns.

For $n \geq 3$, we can partition the set of valid lines into eight disjoint subsets depending on whether or not the line contains a move and on which of the four ending patterns occur. These are shown in Table 1 along with the variables we use to count each type.

Table 1. Variables for the eight disjoint subsets.

	does not contain a move	contains a move
ends in ccd	s_n	w_n
ends in cdc	t_n	x_n
ends in cdd	u_n	y_n
ends in cde	v_n	z_n

The number of lines in each of the subsets can be related through a system of recurrence relations. For example, s_n is the number of valid lines of n candies that do not contain a move and end with a ccd pattern. Such a line is formed by adding a candy to a valid line of length $n - 1$ that does not contain a move and ends in a pair. The added candy can be one of $q - 1$ colors—any color that is different from the final pair. Thus, we see that $s_n = (q - 1)u_{n-1}$.

Similarly, t_n is the number of valid lines of n candies that do not contain a move and end with an cdc pattern. Such a line is formed by adding a candy to a valid line of length $n - 1$ that does not contain a move and ends with two distinctly colored candies. Consider the three possible cases that end with two distinctly colored candies and extend each line by matching the color of the new candy (position n) to the color of the candy in position $n - 2$:

- If a line that ends in $\ldots ccd$ is extended to $\ldots ccdc$, a move becomes possible.
- If a line that ends in $\ldots cdc$ is extended to $\ldots cdcd$, no move will be created.
- If a line that ends in $\ldots cde$ is extended to $\ldots cded$, no move will be created.

Thus we see that $t_n = t_{n-1} + v_{n-1}$.

Similar arguments yield recurrence relations for the remaining sequences. For $n \geq 4$ they are

$$\begin{aligned}
s_n &= (q-1)u_{n-1}, \\
t_n &= t_{n-1} + v_{n-1}, \\
u_n &= s_{n-1} + v_{n-1}, \\
v_n &= (q-2)s_{n-1} + (q-2)t_{n-1} + (q-2)v_{n-1}, \\
w_n &= (q-1)y_{n-1}, \\
x_n &= s_{n-1} + w_{n-1} + x_{n-1} + z_{n-1}, \\
y_n &= t_{n-1} + w_{n-1} + x_{n-1} + z_{n-1}, \\
z_n &= (q-2)w_{n-1} + (q-2)x_{n-1} + (q-2)z_{n-1}.
\end{aligned} \tag{3}$$

The initial values of each sequence can be easily calculated:

$$s_3 = t_3 = u_3 = q(q-1),$$
$$v_3 = q(q-1)(q-2),$$
$$w_3 = x_3 = y_3 = z_3 = 0.$$

Let b_n be the number of valid lines that contain a move. Then

$$b_n = w_n + x_n + y_n + z_n = a_n - (s_n + t_n + u_n + v_n).$$

Table 2 shows the percentage of valid lines that contain a move, for $5 \leq n \leq 10$ candies per line, assuming six possible colors of candy.

Table 2. The percentage of valid lines of length n that contain a move.

n	b_n	a_n	percentage
5	630	7,200	8.75
6	5,220	42,150	12.38
7	39,300	246,750	15.93
8	278,850	1,444,500	19.30
9	1,906,800	8,456,250	22.55
10	12,704,100	49,503,750	25.66

If we assume that each valid line is equiprobable, then in the 5 by 8 grid which begins Level 1 of Candy Crush with six colors, we can expect on average only 0.965 of the 5 rows will contain an in-line move and only 0.7 of the 8 columns will contain an in-line move.

In fact, most of the moves in Candy Crush are not of the in-line variety.

Two lines of candy

If we have a $2 \times n$ grid, then there are $(a_n)^2$ possible configurations of candy placements that will avoid having 3 in a row—any valid configuration of the first line can be paired with any valid configuration of the second line, so the product principle applies. There are many more than $(b_n)^2$ configurations that have moves, however!

In addition to the moves contained within a single line, we must consider the various ways to fill a 2×3 grid so that it contains a move; see Figure 4.

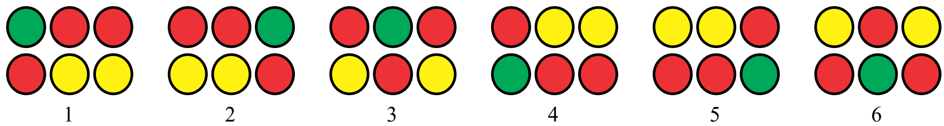

Figure 4. Two adjacent circles can be swapped to obtain three black (red) circles in a row.

In each of the six cases, the circles shown as solid black (red in the color version) can be colored in q ways. The gray (green) circle must be a different color to avoid 3 in a row, so it can be colored in $q - 1$ ways. Each white (yellow) circle can be filled in with any color, as long as both are not black (red). Therefore, the pair of white (yellow) circles can be filled in $q^2 - 1$ ways. Thus, each case can be colored in $q(q-1)(q^2-1)$ ways.

Note there is overlap between cases, however. Figure 5 shows the intersections between pairs that can be colored in $q(q-1)^2$ ways—the black (red) circles can be any of the q colors, and the gray (green) and white (yellow) circles cannot be the same color as the black (red) circles.

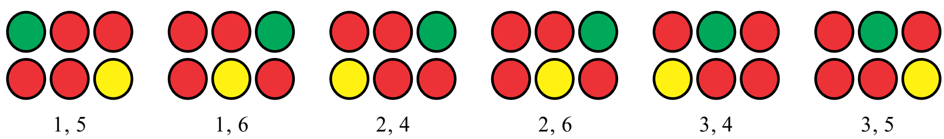

Figure 5. These configurations contain more than one possible move.

Figure 6 shows the intersections between pairs that can be colored in $q(q-1)$ ways.

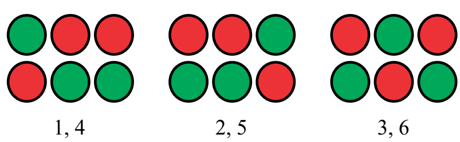

Figure 6. These overlapping configurations use just two colors.

Therefore, the total number of ways to fill in a valid 2×3 grid so that a move is possible is

$$6\left(q(q-1)(q^2-1)\right) - \left(6q(q-1)^2 + 3q(q-1)\right) = 6q^4 - 12q^3 + 3q^2 + 3q.$$

There are $a_3^2 = (q^3 - q)^2$ valid 2×3 grids. When $q = 6$, we see that 5,310 of the 44,100 valid 2×3 grids, or about 12% of them, contain a move.

In theory, the number of $2 \times n$ grids that contain a move can be expressed using a system of recursive equations depending on the end behavior of the last 2×3 grid. However, when $q > 2$, the large number of possible end behaviors makes this approach impractical.

Licorice and coconut jelly beans

Suppose that after a sequence of successful color bombs, only two colors remain. Now (1), the recursive formula for the number of valid lines of candy, becomes

$$a_1 = 2,\ a_2 = 4,\ a_n = a_{n-1} + a_{n-2}.$$

Recall that F_n, the nth Fibonacci number, satisfies the same recurrence relation,

$$F_1 = 1,\ F_2 = 1,\ F_n = F_{n-1} + F_{n-2}.$$

In fact, when $q = 2$, we see that $a_n = 2F_{n+1}$.

Now consider the formula for b_n, the number of valid lines that contain a move. The system of recurrence relations in (3) simplifies to

$$\begin{aligned} s_n = t_n = u_n &= 2, \\ v_n &= 0. \end{aligned}$$

Since $b_n = a_n - (s_n + t_n + u_n + v_n)$, the number of lines that contain a move is given $b_n = a_n - 6 = 2F_{n+1} - 6$. The equality $b_n = a_n - 6$ can be observed directly: If we have 2 colors and need to avoid having three in a row of the same color, then the first three candies in a line can be colored in six different ways. Each of these initial colorings can be uniquely extended to a line of length n that avoids having three consecutive candies of the same color and avoids having a move. The last three candies determine the color of the next candy to be added in order to preserve the property of not containing a move, as shown in Table 3.

Table 3. How the last three candies determine the color of the next candy.

last three candies	next color
BBW	*W*
BWB	*W*
BWW	*B*
WBB	*W*
WBW	*B*
WWB	*B*

With six colors we observed that about 12% of the valid 2×3 grids contain a move. With only two colors, half of the valid 2×3 grids, 18 out of 36, contain a move. As n increases, the percentage of $2 \times n$ grids that contain a move also increases. We show next that, when $n > 5$, only six of the $2 \times n$ grids do not contain a move, so that the number of valid $2 \times n$ grids containing a move is $a_n^2 - 6 = 4F_{n+1}^2 - 6$.

Consider the following labeling of two lines of candy.

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} & \dots \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & \dots \end{pmatrix}$$

There are six valid ways to fill positions c_{11}, c_{12}, and c_{13}. We claim that each choice determines a unique $2 \times n$ grid with no move. As an example, consider the case where $c_{11} = B$, $c_{12} = W$, and $c_{13} = B$. We do not want the grid to contain a move, so the choice of the first three candies determines the top row,

$$\begin{pmatrix} B & W & B & W & B & W & \dots \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & \dots \end{pmatrix}.$$

Note that $c_{22} = W$, else c_{12} and c_{22} could be swapped to match three of a kind. Similarly, the colors c_{23}, c_{24}, and c_{25} are forced, giving

$$\begin{pmatrix} B & W & B & W & B & W & \dots \\ c_{21} & W & B & W & B & c_{26} & \dots \end{pmatrix}.$$

Finally, with $c_{24} = W$, we see that c_{21} must be black. Once the color of the first three candies in a row is determined, the entire row is determined, so this forces a single valid coloring with no move for the entire $2 \times n$ grid.

The other five cases are similar. The choice of c_{11}, c_{12}, and c_{13} uniquely determines a valid grid with no possible move. This proves that there are only six ways to complete a valid $2 \times n$ grid using 2 colors so that no move is possible.

Finally, with two colors, we show that every 3×3 grid contains a move.

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$$

Without loss of generality, let c_{11} be black. First consider the case where c_{12} is the same color. If $c_{12} = B$, then $c_{13} = W$ in order to avoid three in a row and $c_{23} = W$ to avoid a move.

$$\begin{pmatrix} B & B & W \\ c_{21} & c_{22} & W \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$$

Now $c_{33} = B$ to avoid three in a row and $c_{32} = B$ to avoid a move, so $c_{31} = W$ to avoid three in a row.

$$\begin{pmatrix} B & B & W \\ c_{21} & c_{22} & W \\ W & B & B \end{pmatrix}$$

This grid cannot be completed—a black c_{22} would give three in a row while a white c_{22} would make a move possible. Therefore, $c_{11} \neq c_{12}$. By symmetry we also know $c_{11} \neq c_{21}$. Now consider the case when $c_{12} = c_{21} = W$.

$$\begin{pmatrix} B & W & c_{13} \\ W & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$$

Candies c_{13} and c_{31} must both be black to avoid a move.

$$\begin{pmatrix} B & W & B \\ W & c_{22} & c_{23} \\ B & c_{32} & c_{33} \end{pmatrix}$$

This forces $c_{22} = W$ to avoid a move, which forces $c_{23} = c_{32} = B$ to avoid three in a row.

$$\begin{pmatrix} B & W & B \\ W & W & B \\ B & B & c_{33} \end{pmatrix}$$

Again, the grid cannot be completed—a black c_{33} would give three in a row (twice!) while a white c_{33} would make a move possible. Therefore, it is not possible to complete the grid if $c_{11} \neq c_{12}$. This proves that when only two colors are used, every valid 3×3 grid contains a move. Therefore, counting the number of valid $m \times n$ grids that contain a move is the same as counting the number of $m \times n$ grids that avoid having three consecutive candies of the same color.

Further questions to explore

Considering three or more lines of candy becomes quite a bit more complicated, even with the simplifying assumption of only two colors. Counting the number of valid configurations with a move on an $m \times n$ board when m and n are larger than 3 will require new ideas. Here are some other directions to explore.

- In order to advance, players must match candies in more complex patterns. For example, a wrapped candy results from matching candies in a T or L shape, and a striped candy and color bomb result from matching 4 and 5 in a row, respectively. Identify the set of possible configurations that allow a player to obtain one of these patterns on the next move, or in k additional moves.
- In Candy Crush Soda Saga, an additional move is introduced: creating a 2×2 square of the same color results in a jelly fish. How does this impact the number of valid configurations that have moves?

Sweet! Happy crushing.

Summary. In the popular game Candy Crush, differently colored candies are arranged in a grid and a player swaps adjacent candies in order to crush them by lining up three or more of the same color. At the beginning of each game, the grid cannot have three consecutive candies of the same color in a row or column, but it must be possible to swap two adjacent candies in order to get at least three consecutive candies of the same color. How many starting configurations are there? We derive recurrence relations to answer this question for a single line of candy, and also for a pair of lines in the 2-color version of the game.

References

1. King.com Ltd., http://company.king.com/our-games/, 2015.

